Quality Control Standards
- A Laboratory Perspective

The routine analysis of reference materials is an integral part of best practice for all analytical laboratories. The data generated from these standards is used to ensure that the analytical processes are ‘in control’.

Acceptable practice is for exploration companies to monitor a laboratory’s performance by including their own standards among the samples that they submit for analysis. This enables them to independently assess a laboratory’s data to determine accuracy, precision and bias.

The analytical data generated from a reference material for a particular element should be symmetrically distributed around the assigned value. This distribution is defined by the mean and the standard deviation of the data. Limits are set within which the analytical results may be expected to lie while the process is in a state of statistical control. Statistically, 99.7% of all data generated for an individual element in a particular standard must fall within the ‘3-sigma’ limits if the analytical process is ‘in control’.

It is now commonplace for customers to specify that the data returned from their standards must fall within the ‘2-sigma’ limits of the normal distribution. Statistically we can expect 95% of all acceptable data to fall within these limits and 5% to fall outside. The question that must be addressed is what data has been used to set these limits.

Why have the ‘2-sigma’ limits become the norm?

This may well have come about because certification programs are subject to more rigorous statistics than were previously used. Many standards now quote the mean value obtained during the certification process and the calculated 95% confidence interval within which the real value is expected to lie. The statistics that accompany certified reference materials are a measure of the homogeneity of the material and the agreement between the laboratories that participated in the certification exercise.

What exactly is a confidence interval?

Quantitative analysis provides an estimate of the value being sought and always involves some level of uncertainty. The certified value of a standard is the best estimate of the concentration available from the certification data. The confidence interval defines the range of values, calculated from the mean and standard deviation, which is expected to include the population mean with a stated level of confidence. The 95% confidence interval quoted denotes that if the certification program were to be repeated many times, the certified mean value would be expected to fall within the defined limits 95 times out of every 100 programs.
How may a geologist or metallurgist utilise their own standards to monitor a laboratory’s performance?

If the certification data quotes a between-laboratories standard deviation (usually denoted as $s_r$), control limits may be set at $\pm 2s_r$ from the quoted mean. This will provide an initial estimate of what a laboratory may be able to achieve. Alternatively, a laboratory will usually quote a level of precision that can be applied to individual methods. These values are based on evaluation of data that has been generated over a long period using different operators, equipment and calibration standards. A typical value for routine fire assay gold precision is 10%, so this would make a good starting point for setting control limits. Experience may dictate that control limits can be tightened slightly, but they will never approach the 95% confidence intervals quoted for a reference material. In the long run it is impossible to obtain more precise data on a standard than what the analytical method is capable of achieving.

Does the quoted laboratory precision level apply at all concentrations above the limit of detection of a method?

The short answer is no. A laboratory will quote a minimum value above which the expected precision will apply – usually 50 times the limit of detection of a method. Why is this justified? There is always some measurement uncertainty associated with any determination. A measured value becomes believable when it is larger than the uncertainty associated with it. This point is called the limit of detection and is arbitrarily defined as 3 times the standard deviation of multiple determinations on a reagent blank ($3s_0$). The lower level where measurements become meaningful is called the limit of quantitation (LOQ), which is usually defined as $LOQ = 10s_0$. For routine fire assay gold determinations the relative confidence in a measured value is around $\pm 100\%$ at the 95% probability level. This is equivalent to $\pm 0.03$ ppm at the 0.03 ppm gold concentration. This is the minimum measurement uncertainty present in any gold determination. The method will not allow us to achieve better precision over the long term than this value.

Keeping the mean value of a standard within the control limits (either 2-sigma or 3-sigma) does not necessarily mean that the analytical process is ‘in control’. The distribution of the data within these limits is also important. Having all data sitting above or below the allocated mean value indicates that the laboratory has a bias with this element.

It must be recognised that the certified value assigned to a standard is the best estimate available from the data gathered. The confidence interval quoted states a range in which the real value may exist. One would expect the true value to lie closer to the mean value than to one or other of the limits quoted. A laboratory that generates a long-term mean value within $\pm 50\%$ of the stated confidence interval should not be considered to have a bias if their data is normally distributed about their mean value.

A laboratory’s QC data may show a periodicity when plotted. This may well be related to the introduction of new instrument calibration solutions. Good laboratory practice dictates that new calibration solutions must be checked against the old. These solutions are never exactly the same so it follows that there will be a slight variation in the data.
Is it possible for laboratory and company standards to provide conflicting points of view?

This can happen if any of the following situations arise:

- Certain producers of laboratory standards grind and screen their materials very finely to ensure that the resulting standard is quite homogeneous. This is particularly applicable to gold standards where it is essential to ensure that no ‘coarse’ gold particles are present. These materials may be difficult to mix mechanically with the fire assay flux, leading to ‘balling’ and incomplete fusion giving a low bias. Hand mixing of these standards will provide the laboratory with better precision and accuracy.

- The majority of fire assay gold determinations undertaken worldwide use a 30 gram sample weight. This would be reflected in most gold certification exercises. A critical factor in fire assay gold recovery is the flux : sample ratio. A minimum ratio of 4:1 is recommended to ensure sample dissolution and recovery of the gold. If this is not achieved one would expect to find a low gold recovery. A more complex sample matrix will require more flux to achieve optimum gold recovery.

- The higher the sample weight taken for analysis, the more difficult it can be to achieve a satisfactory mix of sample and flux particles. The optimum requirement is that each sample particle is completely surrounded by flux particles. If this is not achieved, sample dissolution may be compromised and low gold recoveries obtained.

- There is a marked difference between sample matrices, leading to less than ideal element recoveries. This may occur if the customer submits gold standards with a sulfide matrix certified by fire assay for an aqua regia or cyanide extractable gold determination. Low recoveries will usually result unless the material is pre-roasted.

- It is always necessary to compare ‘apples with apples’. Base metal standards are usually certified using a ‘total digest’ procedure. Using a partial extraction procedure (i.e. an aqua regia digestion) will result in low recoveries for elements that are present in minerals that are incompletely digested. A similar problem can occur if trying to compare the analytical precisions of ore-grade and geochemical methods on a single control chart.

- It is meaningless to submit an ore grade standard with a consignment of exploration trace element samples, and vice-versa. The QC data received will not allow any meaningful assessment of the analytical data in the attached report. It may even increase the total cost if the laboratory undertakes an ore grade analysis on the standard.

- Standards with a sulfide matrix have a limited shelf life once they are exposed to the atmosphere. Values slowly reduce over time due to oxidation. This effect is more obvious with base metal standards due to higher elemental concentrations usually encountered. Nickel standards are especially prone to degradation as they often contain elevated concentrations of pyrrhotite, which is very reactive.

- Most standards are certified on a dry weight basis. Certain sample matrices are hygroscopic (e.g. nickel laterite), absorbing a considerable amount of moisture upon standing. If the laboratory fails to dry these materials prior to weighing, low and variable recoveries will occur.

- A blank or a low level base metal standard submitted immediately after a highly mineralised sample will usually show elevated values due to carryover from the previous test solution. ICPAES or ICPMS operating protocols allow a standard wash-out time between test solutions, but this may not be sufficient to eliminate carry-over from more concentrated solutions.

Sending randomly selected samples to a second laboratory for check analysis is sound quality practice. It is recommended that this be done on a continual basis as it allows potential problems to be detected early, rather than having a major issue at the end of a program.

References

Quality Assurance of Chemical Measurements

Using the Correct Control Limits
Lynda Bloom, Maureen Leaver

Association of Exploration Geochemists Newsletter, No. 115